一种资源发现模型的形式化表示及其应用

包义保，贺龙涛，时金桥，方滨兴，王丽宏

(中国科学院计算技术研究所信息智能与信息安全研究中心，北京 100190)

摘 要： 类似于P2P的动态服务系统的资源发现策略、评介指标及资源发现成本之间的关系对于资源发现系统的优化非常重要。本文以资源发现为目的对服务系统进行建模，给出其一般化的表示方法，在此基础上，作为特例研究了其特化模型中枚举策略条件下资源发现成本的表示方法和估算公式，并就枚举策略的优化过程进行了详细的讨论。

关键字： 网络测量 资源发现 资源枚举

A Formalized Model and Application of Resource Discovery

Yibao BAO, Longtao HE, Jinqiao SHI, Bingxing FANG, Lihong WANG

Research Center of Information Intelligence and Information Security
Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190

Abstract: In dynamical service systems, such as P2P, relationship between resource discovery strategy, evaluation and consumed resource for this is essential for the optimization of the resource discovery systme. In this paper, a formalized resource discovery model of dynamic network is given. Then, as an example, we research it in its specialized model scene with an enumeration strategy and give the mathematic formulas of the consumed resource in different situations, and we also discuss the optimizing method of enumeration strategies in detail.

Key words: Network measure, Resource discovery, Resource enumeration.

1. 引言

随着世界范围内 Internet 的快速发展，网络体系结构取得了一些重大进展，类似于覆盖网络(Overlay Network)的自治网络体系架构变得越来越流行，如 P2P[0]、·TOR[0](一种匿名通信网络架构)等。这些“自组织”类型服务系统的共同特点是系统中的结点群不断变化，每个实体在任何时刻都可能自主地进入或离开网络系统。

虽然自组织服务系统具备良好的适应能力，为人们提供了丰富的服务功能，但同时类似于 P2P 之类的网络服务系统时常充斥着大量的非法内容。由于这些构建于标准网络之上的覆盖网络结构灵活，管理松散，常使用自定义（或非标准）的协议，具有较强的不确定性和难管理性，给网络管理造成了严重的负担。

为了维护网络空间的正常秩序，对网络服务系统进行更好的管理，就需要建立专门的资源发现系统，通过某种资源发现策略对参与此类服务系统的结点群进行测量。同时，由于计算资源和网络资源（如带宽等）的紧缺，建造专用资源发现系统不仅需要考虑投入规模问题，还需要考虑投入优化问题，这可通过建立资源发现策略和资源发现成本（即资源发现系统通过某种发现策略完成测量任务所需要投入的资源，如计算资源和网络带宽资源等)之间的关系，从而通过调整资源发现策略的相关参数，降低资源发现成本，达到系统管理和效益优化的双重目的。本文将就这一问题进行研究：第 2 小节建立了通用的资源发现与分析模型，在此基础上，第 3 小节研究了特化模型中特定资源发现策略(即资源发现系统通过枚举操作获取服务系统中服务结点的策略，简称为枚举策略)、评价指标和资源发现成本之间的关系及资源发现成本的估算公式，并就枚举策略的优化进行了详细的讨论。

基金项目：国家重点基础研究发展计划（“973”计划）基金资助项目(2004CB318109 & 2007CB311100)

作者简介：包义保（1976.8—），男，安徽无为人，讲师，博士生， baoyibao@software.ict.ac.cn;
方滨兴（1960.7—），男，江西万年人，中国工程院院士，教授，博士生导师。

2. 模型的建立

本文所研究的服务系统是一个瞬时系统，其结构、规模、提供服务的结点数目等随时都有可能发生变化，一般来说，在任一 Δt 时间段内，新入和退出的服务结点的数目都是随机的，也就是说，结点进入和退出服务系统的过程为一随机过程[0]，从而使得任一时刻服务系统中活跃结点的数目也是随机的，如图 1 所示，其中，资源发现系统所感兴趣的实体称为资源，每个资源实体称为结点，服务系统中能够提供服务功能的结点称为活跃结点，所有活跃结点及其管理策略统称为资源池，资源池中的活跃结点是动态变化的。

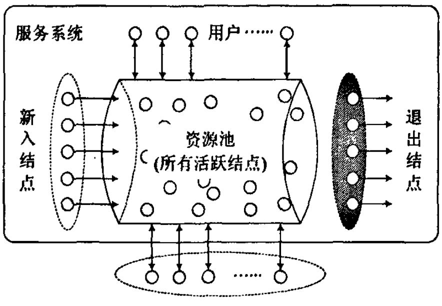


图 1　带有资源发现功能的动态服务系统工作原理：资源池中的结点是动态变化的。

为了发现资源池中的结点，资源发现系统需要根据资源池的实际情况 Q，在一定的资源发现成本 C 及一定的评价指标 R 下，采用相应的资源发现策略 S，从而完成资源发现任务，则上述系统可形式化地表示成 $Model=(Q,R,C,S)$，其目标就是在达到评价指标 R 的前提条件下，通过调整资源发现策略 S 使得资源发现成本 C 最小化，或通过制定合理的评价指标 R 和资源发现策略 S 而使得资源发现成本 C 处于可承受的范围内。

将上述 4 元组所表示的模型映射到具体的服务系统时，每个元素都需要根据其自身的特点灵活选定。例如 Q 主要涉及到新入结点、活跃结点及退出结点的分布情况和统计特性，此外，还可能涉及其它因素，如资源发布策略等。资源发现成本 C 不仅取决于资源池的实际情况 Q，还和资源发现策略 S、评价指标 R 有着重要的关系，一般来说，可通过它们之间的这种关系，建立数学方程，从而达到优化 C 的目的。资源发现策略 S 受认知水平及现实环境的影响特别大，需要具体问题具体分析，但也存在一些非常特殊的资源发现策略，如下文所研究的资源发现系统通过模拟服务系统的用户来枚举资源池中结点的策略，这一策略形象地称为枚举策略。

3. 模型的分析

$Model=(Q,R,C,S)$ 是一个形式化的模型，应用到具体的资源发现系统时需要进行实例化处理。下面将通过实例说明如何将 $Model=(Q,R,C,S)$ 模型用于分析具体的资源发现系统，为此，需要对模型做出必要的特化处理。

3.1 模型的特化

资源发现系统拟通过简单的枚举操作获取资源池中结点的相关信息，这是最简单、最常用的一类资源发现策略，称为**枚举策略**。此时，资源发现系统也形象地称为资源**枚举系统**，其工作过程可抽象表示为：

（1）枚举系统模拟服务系统的用户从资源池中获取结点的相关信息(如 IP 地址、数字证书等)，称为枚举请求；

（2）枚举系统可以同时并行发出多个枚举请求，并且假设服务系统无法感知这些枚举请求来自于同一个结点。每个枚举请求称为一个**枚举线程**，枚举系统并行运行枚举线程的数量称为**枚举并行度**，用 r 表示。

（3）枚举线程可能仅枚举一次就终止，也可能迭代运行。枚举线程两次枚举之间的最小时间时

隔为 T_e；无论具体情况如何，枚举线程从发出枚举请求到获得服务节点的这一过程称为枚举线程的一次**枚举操作**。

（4）当枚举线程迭代运行时，每次枚举都是独立的。单位时间内枚举线程迭代运行次数称为**迭代频率**，用 w 表示，当 w 为 0 时，表示枚举线程不进行迭代。

（5）在实际应用中，考虑到资源投入量的限制、费用等问题，一般不需要枚举出资源池中所有的服务结点，而仅需要"平均枚举出一定比例"的结点即可，这一比例称为**枚举测度**。例如，对于一些服务系统，如果能够发现其中 80% 以上的结点，则整个服务系统就可处于安全的可控状态。系统中需要关注 4 种类型的枚举测度，如表 1 所示。

表 1　服务系统资源枚举模型中 4 种类型的枚举测度

符号表示	意义
P_{evl}	枚举操作后系统期望达到的枚举测度
P_{deg}	枚举操作后系统实际达到的枚举测度
P_0	系统初始的枚举测度，可为任意数值
P_{min}	系统所能忍受的最小枚举测度
应该满足：$1 \ge P_{deg} \ge P_{evl} \ge P_{min} \ge 0,\ 1 \ge P_0 \ge 0$	

本文研究的目标是要建立资源发现成本的表示方法和估算公式，从而为优化资源发现策略提供依据。在枚举策略条件下，枚举系统的资源发现成本 C 称为**枚举成本**，它是一个需要形式化的参量，需要反映计算量、通信带宽等多方面的限制，可使用 3 元组 $C=(r,w,r\cdot w)$ 来进行度量，其中 r 和 w 反映了计算资源总量，而 $r\cdot w$ 反映了网络带宽资源总量。

在枚举策略条件下，枚举成本 C 与资源池中结点的更新速度、数量及服务系统资源发布策略有着密切的关系。为此，令 $U_t(\Delta t)$ 表示从 t 时刻起在 Δt 时间段内新入结点的数量，则 $U_t(\Delta t)$ 为一随机变量，其相应的数学期望记为 $u(\Delta t)=E(U_t(\Delta t))$；令 $O_t(\Delta t)$ 表示从 t 时刻起在 Δt 时间段内退出结点的数量，则 $O_t(\Delta t)$ 为一随机变量，其相应的数学期望记为 $o(\Delta t)$；令 $M(t)$ 表示时刻 t 服务系统中存活结点的数量，则 $M(t)$ 也为一随机变量，其相应的数学期望记为 $m=E(M(t))$。$u(\Delta t)$ 和 m 都可通过统计分析实验数据得出，或在大量实验数据分析的基础上做出合理的假设。此外，枚举过程和服务系统的资源发布策略 D 有关：有的服务系统一次仅从其所拥有的结点全集或某个子集中按照某种策略选择若干个服务结点发布到请求服务的用户，而有的服务系统一次性向请求服务的用户发布其拥有的全部服务结点。由于资源发布策略的巨大差异性，因而无法忽略其对枚举成本的影响。因此，在枚举策略条件下资源池实际情况 Q 可表示为 $Q=(M(t),U_t(\Delta t),O_t(\Delta t),D)$，如果仅关心其统计特征，则 Q 也可表示为 $Q=(m,u(\Delta t),o(\Delta t),D)$。

为了进一步简化资源池实际情况 Q 的复杂性，可作如下处理：

(1) 设服务系统的资源池容纳服务结点的能力是无限的，如果把新入结点看成排队等候服务的顾客，把其在资源池中处于活跃状态的过程看成接受服务的过程，把其退出资源池的过程看成完成服务的过程，则整个服务系统就是一个随到随服务、具有无限服务窗口的服务系统[5]。本文仅考察服务系统趋于稳定状态后的情形，此时，新入结点和退出结点保持平衡[5]，且资源池中结点数目的变化也比较均匀，从而可通过统计特征来刻画其实际情况。当服务系统达到稳定状态后，$m=E(M(t))$ 为一确定值，单位时间内新入结点的平均数目和退出结点的平均数目相同，记为 λ[1][2]，则 $u(\Delta t)=o(\Delta t)=\lambda\cdot\Delta t$，因而 Q 可简化表示为 $Q=(m,u(\Delta t)=\lambda\cdot\Delta t,D)$。

(2) 设每个枚举线程每次枚举所获得的节点都是由服务系统随机选取、且在资源池中都是独立、均匀、同分布的，其平均数目为 $N_e(0<N_e\le m)$。另外，用户连续两次枚举请求之间的时间间隔至少为 T_d，且 $T_e\ge T_d$，因而，资源发布策略可表示为 $D=(N_e,T_d)$。

至此，一个简化的带有资源枚举功能的服务系统可表示为 $Model=(Q,R,C,S)$，其中资源池实际情况 $Q=(m,u(\Delta t)=\lambda\cdot\Delta t,D=(N_e,T_d))$，枚举策略 $S=(r,w,P_0)$，评价指标 $R=(P_{deg}\ge P_{evl}\ge P_{min},1/w\ge T_e,T_e\ge T_d,\cdots)$，枚举成本 $C=(r,w,r\cdot w)$。下面研究该特化模型中各种枚举策略条件下枚举成本 C 的估算方法及枚举策略的优化。

3.2 无迭代、无记忆资源枚举策略

设资源池实际情况 $Q=\left(m,0,D=(N_e,T_d)\right)$，枚举策略 $S=(r,0,0)$，则其相应的枚举成本为 $C=(r,0,0)$。由于该策略下，枚举线程无迭代，且初始枚举测度为 0，因而称之为“无迭代、无记忆资源枚举策略”。该枚举策略事实上反映了模型在静态条件下的一些内在规律，此时，枚举过程等价于实验 1。

实验 1：服务系统（记为 DY1）达到平衡状态后，平均有 m 个结点。r 个枚举线程并行地从 DY1 中随机等概地获取这些结点的相关信息。不防假设枚举操作过程中，系统中的资源池保持不变。那么这 r 个枚举线程运行完毕后，服务系统的枚举测度为多少？

在实验一中，对于任一线程来说，任一结点没有被枚举到的概率为 $\frac{m-N_e}{m}=1-\frac{N_e}{m}$，那么对 r 个并行独立运行的线程来说，任一个结点没有被枚举到的概率为 $\left(1-\frac{N_e}{m}\right)^r$。实验 1 完成后，$m$ 个结点中平均有 $m\cdot\left(1-\frac{N_e}{m}\right)^r$ 个结点没有被枚举过。因此，对于由图 1 所表示的资源发现模型，枚举并行度 r 和枚举测度 $P_{\deg}$、P_{evl} 之间的关系可表示为

$$P_{\deg}=\frac{\left(m-m\cdot\left(1-\frac{N_e}{m}\right)^r\right)}{m}=1-\left(1-\frac{N_e}{m}\right)^r\geq P_{evl}\text{，令 } s=\frac{N_e}{m},0<s\leq 1$$

则有

$$P_{\deg}=1-(1-s)^r\geq P_{evl} \qquad \text{（公式 1）}$$

当 $s=1$ 时，即 $N_e=m$，记 $0^0=1$，则 $r_{min}=1$。当 $0<s<1$ 时，即 $N_e<m$，P_{deg} 是关于 r 的增函数，则一定存在最小的正整数

$$r_1=r_{\min}=\left\lceil\frac{\log(1-P_{evl})}{\log(1-s)}\right\rceil \qquad \text{（公式 2）}$$

使得上述不等式成立。例如，当 $s=1/2$、$P_{evl}=75\%$ 时，可推导出 $r\geq 2$，即 $r_{\min}=2$。图 2 反映了 s、r、P_{deg} 和 P_{evl} 之间的关系：P_{deg} 随着 r 的增加和 s 的减小而增加。

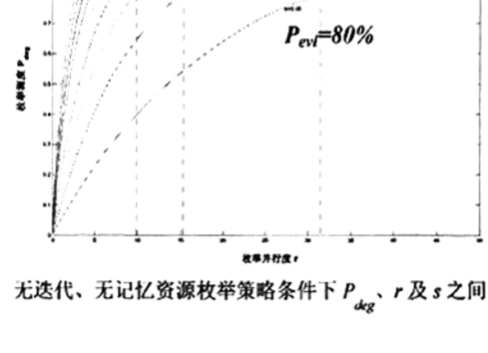


图 2　无迭代、无记忆资源枚举策略条件下 P_{deg}、r 及 s 之间的关系

由上述分析可知，当 s 非常小时，r_{min}(即枚举系统的并行度)将是一个非常大的值，在实际的运行

环境中，r_{min}过大是不现实的，那么如何降低系统的枚举并行度呢？此外，由于实际系统中的资源池是动态变化的，从而使得系统的枚举测度一直呈下降趋势，那么，如何通过枚举操作补偿和保持系统的枚举测度呢？要解决以上两个问题，仅通过本节所分析的枚举策略是不够的，下面将讨论两种更加复杂的资源枚举策略。

3.3 无迭代、有记忆资源枚举策略

设资源池实际情况$Q=(m,u(\Delta t),D=(N_e,T_d))$，枚举策略$S=(r,0,P_0)$，其相应的枚举成本$C=(r,0,0)$。该枚举策略将部分反映模型在动态条件下系统枚举测度变化的内在规律，此时，枚举过程等价于下述实验2。

实验2：服务系统（记为DY2）达到平衡状态后，平均有m个结点。在某一初始时刻t，系统中每个结点被枚举过的概率为P_0。由于DY2中的结点是动态变化的，Δt时间后，有$u(\Delta t)$个结点退出系统，同时有同样多的结点进入系统。此时，有r个枚举线程并行地从DY2中随机等概地获取这些结点的相关信息。不防假设在枚举操作过程中，资源池保持不变。那么这r个枚举线程运行完毕后，服务系统的枚举测度是多少？

在实验2中，枚举系统进行枚举动作之前的那一时刻记为$t'=t+\Delta t$，DY2中的结点集可分成两个子集合：一个子集是DY2中原有的、没有退出的结点集，记为A，则$E(|A|)=m-u(\Delta t)$，A中的每个结点被枚举过的概率为P_0；另一子集为新加入的结点集，记为B，则$E(|B|)=u(\Delta t)$，B中每个结点被枚举过的概率为0。

令X表示t'时刻DY2中被枚举过的结点总数，X_1表示t'时刻A中被枚举过的结点总数，X_2表示t'时刻B中被枚举的结点总数，则有$X=X_1+X_2$，即离散随机变量X和X_1、X_2之间存在线性关系，因而其数学期望也存在相应的线性关系[0]，如下式所示。

$$E(X)=E(X_1)+E(X_2)=E(|A|)\cdot P_0+E(|B|)\cdot 0=(m-u(\Delta t))\cdot P_0$$

令Y表示t'时刻DY2中没有被枚举过的结点总数，则$E(Y)=E(\overline{X})=m-(m-u(\Delta t))\cdot P_0$。

在t'时刻之后，枚举系统进行枚举动作(根据假设，枚举动作期间DY2中的结点集保持不变)。设a_i表示i个并行线程枚举后DY2中没有被枚举过的结点数目的数学期望，则

$$a_0=E(Y)=E(\overline{X})=m-(m-u(\Delta t))\cdot P_0$$

$$a_i=a_0\left(1-\frac{N_e}{m}\right)^i$$

因而DY2在Δt时间段后，重新启动一次并行度为r的枚举操作后，系统的枚举测度可提高到

$P_{\deg}=1-\frac{a_r}{m}=1-\left(1-\left(1-\frac{u(\Delta t)}{m}\right)\cdot P_0\right)\cdot\left(1-\frac{N_e}{m}\right)^r$。令$v=\frac{u(\Delta t)}{m}$，$s=\frac{N_e}{m}$，并且要使得枚举后系统的枚举测度$P_{\deg}\ge P_{evl}$，则有

$$P_{\deg}=1-(1-(1-v)\cdot P_0)\cdot(1-s)^r\ge P_{evl}，\ s,v\le 1。\qquad（公式3）$$

$$r_2=r_{\min}=\left\lceil \log_{(1-s)}\left(\frac{1-P_{evl}}{1-P_0+v\cdot P_0}\right)\right\rceil，\ 0<s<1。\qquad（公式4）$$

当$P_0=0$时，（公式3）就退化成（公式1），即DY1是DY2的特例。当$v=1$时，即在Δt时间段内DY2中所有原有结点全部被换出，（公式3）也退化成（公式1），同样，也可以推导出DY1是DY2的特例。当v=0时，即在Δt时间段内DY2中无结点退出，由（公式3）的左半部分可以看出

$P_{\deg} \geq 1-(1-P_0)=P_0$，因此，不需要任何枚举操作就可保持 DY2 的枚举测度不变。

下面讨论 r_2 和 s、v 之间的关系。在 v 一定的情况下，r_2 是关于 s 的减函数。在 s 一定的情况下，r_2 是关于 v 的增函数。因此，对于 DY2 的枚举系统来说，为了降低枚举并行度，在 Δt 一定的条件下，通过增大 s、减小 v 即可。比较（公式 4）和（公式 2）可知，在 $P_0 \neq 0$、其它条件都相同的情况下，$r_2 < r_1$。图 3 反映了 P_{deg} 和 s，v，r，P_0，P_{evl} 之间的关系。

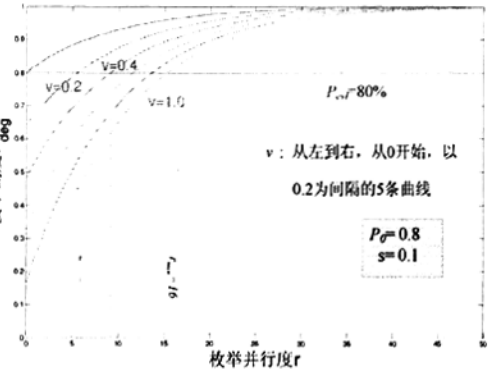


图 3　DY2 枚举策略条件下模型的分析

由 DY2 的分析可知，当一个服务系统的枚举测度降低到一定值后，可以重新启动一次枚举操作以补偿枚举测度的损失，那么两次枚举操作之间的时间间隔 Δt 应该取多少才合适呢？

由 $u(\Delta t)=\lambda\cdot\Delta t$ 可知，v 和 Δt 成正比。因此，为了以最小并行度进行枚举，Δt 越小越好，但枚举线程完成一次枚举操作所需要的时间至少为 T_e，因此有 $\Delta t \geq T_e$。此外，枚举系统的枚举操作是为一定目的服务的，一般设定有最低可容忍枚举测度 $P_{\min}$。在 DY2 中，系统枚举测度最小值发生在枚举操作之前的那一该，其值为 $\frac{(m-u(\Delta t))\cdot P_{evl}}{m} \geq P_{\min}$，即 $\frac{m-\lambda\cdot\Delta t}{m}P_{evl} \geq P_{\min}$，因此有 $T_e \leq \Delta t \leq \frac{m}{\lambda}(1-\frac{P_{\min}}{P_{evl}})$，即

$$\frac{\lambda}{m}\cdot\frac{P_{evl}}{P_{evl}-P_{\min}} \leq w \leq \frac{1}{T_e}$$

由于两次枚举操作之间的时间间隔 Δt 越小，则 r 越小，但是 w 却越大，从而使得 $C=(r,w,r\cdot w)$ 并不一定最优。为此，需要继续研究 $w \neq 0$ 时资源枚举系统的规律。

3.4 有迭代、有记忆资源枚举策略

当 $w \neq 0$ 时，本文仅考虑下述枚举策略：$Q=\left(m,u(\Delta t)=\lambda\cdot\Delta t,D=(N_e,T_d)\right)$，$S=(r,w,P_0=P_{evl})$，其相应的枚举成本 $C=(r,w,r\cdot w)$。该枚举策略事实上反映了服务系统通过迭代枚举操作维持系统枚举测度的内在规律，此时，枚举过程等价于下述实验 3。

实验 3：服务系统（记为 DY3）在实验二的基础上，枚举系统以 $w \neq 0$ 的频率进行迭代枚举操作，且每次枚举操作完毕后，系统的枚举测度都至少提高到初始值 $P_0=P_{evl}$ 的水平，那么枚举并行度 r 和枚举频率之间应满足什么样的关系，使得评价指标 $C=(r,w,r\cdot w)$ 达到优化？

设 Δt 时间段内 DY3 的枚举系统迭代枚举 $n=w\cdot\Delta t$ 轮，q_i 表示第 i 轮枚举操作完成后 DY3 的枚举测度，则 $q_0=P_{evl}$，$q_n \geq q_0$。和 DY2 的分析过程类似，对于 DY3 同样可以近似得出如下等式。

$$q_i=1-\left(1-(1-v)\cdot q_{i-1}\right)\cdot(1-s)^r，其中 v=\frac{u(1/w)}{m}=\frac{\lambda}{m}\cdot\frac{1}{w}，s=\frac{N_e}{m}$$

因而当 $s,v<1$ 时

$$q_i - q_{i-1} = (q_{i-1} - q_{i-2}) \cdot (1-v) \cdot (1-s)^r$$

$$\Rightarrow q_n - q_0 = (q_1 - q_0) \cdot \frac{1 - \left((1-v)\cdot(1-s)^r\right)^{w\cdot\Delta t}}{1-(1-v)\cdot(1-s)^r}$$

因此，要保证 $q_n \ge q_0$，则 $q_1 \ge q_0$ 成立即可，即 $q_1 = 1 - \left(1 - (1 - \frac{\lambda}{m} \cdot \frac{1}{w}) \cdot q_0\right) \cdot (1-s)^r \ge q_0$

令 $x = \lambda / m$，则上式可以写成

$$P_{\deg} = 1 - (1 - (1 - \frac{x}{w}) \cdot P_{evl})(1-s)^r \ge P_{evl} \qquad \text{（公式 5）}$$

（1）当 $x = w$ 时，即在 $1/w$ 时间段内 DY3 中所有原有结点全部被换出，DY3 退化成 DY1。当 $x = 0$ 时，即在 $1/w$ 时间段内 DY3 中无结点退出，$P \ge 1 - (1 - P_{evl}) = P_{evl}$ 恒成立，因此，DY3 不需要任何枚举操作就可保持其枚举测度不变。

（2）图 4 反映了在不同的 x、s 及 $P_{evl}=0.8$ 情况下，P_{deg} 和 w、r 之间的关系。

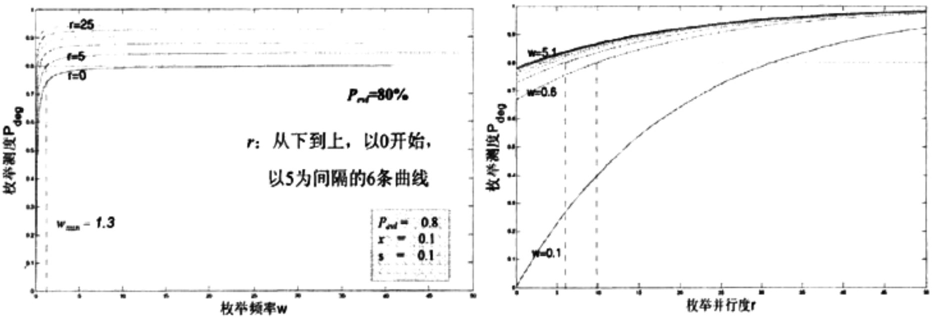
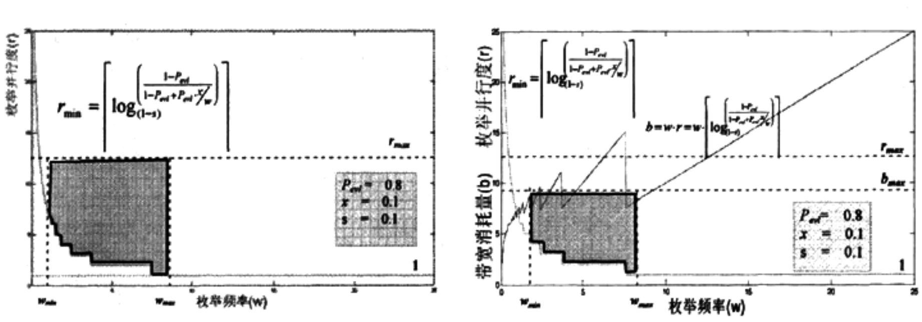


图 4　DY3 枚举策略条件下模型的分析

（3）当 $P_{deg}= P_{evl}$ 时 r 取最小会值 r_{min}：$r_3 = r_{\min} = \left\lceil \log_{(1-s)}\left(\frac{1-P_{evl}}{1-P_{evl}+P_{evl}\cdot x/w}\right)\right\rceil$。图 5(a) 反映了在满足各种决策关系下 r_{min} 的取值范围（阴影部分）。

（4）考虑有带宽限制条件下的 r，w 选取方法，带宽定义为 $b= r \cdot w$，它表示单位时间内的网络流量。图 5(b) 反映了在满足网络带宽约束条件下 r_{min} 的取值范围。



(a) 无带宽限制　　　　(b) 有带宽限制

图 5　满足约束条件下 r 的取值范围

（5）类似于 DY2 的分析，同样可以推导出 $w_{\min} = \frac{P_{evl}}{P_{evl} - P_{\min}} \cdot x \le w \le \frac{1}{T_e} = w_{\max}$。

4. 结论和下一步工作

本文将带有枚举功能的服务系统形式化地表示为 $Model=(Q,R,C,S)$，并针对特化模型中三种不同的资源枚举策略 DY1、DY2 和 DY3 进行了详细的分析，给出了相应的资源发现成本估算公式。DY1 是 DY2 的特例，从某种意义上来说，DY2 也是 DY3 的特例。DY1 适用于那些静态网络的资源枚举问题；DY2 适应于那些定期补偿枚举测度、枚举频率极其低下的资源枚举问题，由于枚举线程的迭代频率极低，由迭代所造成的网络通信量仅是暂时的，不足以对网络造成持续的压力，因而枚举系统的网络带宽资源投入量并不受特别的关注，仅关注其计算资源(即枚举并行度)投入量即可；DY3 反映了实际网络资源枚举系统所需要的计算和网络资源投入量的问题。DY3 在保证 DY2 所实现的全部功能外，通过合理的迭代频率，在保证适度的枚举并行度的同时，也使得因枚举线程的迭代运算所造成的网络通信量不会急剧增加，从而保证两方面都达到优化。

由于服务系统自身的复杂性，针对不同的资源发布策略需要区别对待，本文仅研究了资源发布系统以随机等概的策略发布服务结点的情况。但是，目前已经出现了多种资源发布策略，如以非等概方式发布服务结点，将不同的用户群映射到不同服务结点子集等，一些专用的动态服务系统已经采用更加复杂的资源发布策略来分发服务结点，使得本文所使用的枚举系统可能很难正常工作。此外，本文仅通过枚举测度作为衡量指标来估算枚举系统的资源投入量，在一些应用中，对于何时能够枚举到服务结点的全集也是非常关心的，关于这方面工作的纯数学分析，在文献[00000]中都进行了详细的讨论，如何应用到本文所提出的模型，还需要进一步的研究。

此外，本文给出的模型在规范资源发现系统的设计时虽然能够起到一定的作用，但对于将模型应用于具体的资源发现系统，还需要做大量的工作，如何简化这样的工作，甚至提供更加易用、完善的的模型依然是我们下一步需要做的工作。

参考文献：

[1]Joachim H. Ahrens, Ulrich Dieter (1982). "Computer Generation of Poisson Deviates". ACM Transactions on Mathematical Software 8 (2): 163--179.

[2]Ronald J. Evans, J. Boersma, N. M. Blachman, A. A. Jagers (1988). "The Entropy of a Poisson Distribution: Problem 87-6". SIAM Review 30 (2): 314--317.

[3]Michael Mitzenmacher, Eli Upfal. Probability and Computing: Randomized Algorithm and Probabilistic Analysis. Cambridge University Press. ISBN 978-7-111-20805-1

[4]Roger Dingledine et al. Tor: The Second-Generation Onion Router. Proceedings of the 13th USENIX Security Symposium, ACM, Pp.303-320.

[5]陆传赉.排队论. 京邮电学院出版社.1994 年 5 月第 1 版.ISBN 7-5635-0174-6/O.8.

[6]Wang Qing-Bo et al. Barnet-An Infrastructure for Attribute Addressable P2P Network. Journal of Software vol.14, No.1.

[7]John E.Kobza et al. A Survey of the Coupon Collector's Problem with Random Sample Sizes. Methodol Comput Appl Probab (2007) 9:573–584, DOI 0.1007/s11009-006-9013-3

[8]Adler and S. M. Ross, "The coupon subset collection problem," Journal of Applied Probability vol. 38 pp. 737–746, 2001.

[9]Bosch, "Optimal card-collecting strategies for magic: The gathering," The College Mathematics Journal vol. 31(1) pp. 15–21, 2000.

[10]Stadje, "The collector's problem with group drawings," Advances in Applied Probability vol. 22 pp. 866–882, 1990.

[11]Ivchenko, "How many samples does it take to see all the balls in an urn?" Mathematical >Notes vol. 64 (1) pp. 49–54, 1998.